STARS-c, A Google Earth Engine Tool to Evaluate Long-Term Water Quality Trends Globally

Anna Cardall, Kaylee Tanner, and Gustavious Williams
(Brigham Young University, Utah, United States)

Abstract

We present the SpatioTemporal Analysis of Remotely Sensed chl-*a* tool (STARS-c) tool, based on Google Earth Engine (GEE) and Landsat data. STARS-c characterizes water quality trends for any water body globally. STARS-c allows a user to outline a water body, select a time range, and apply a general or user-provided model to estimate chlorophyl-a (chl-*a*) concentrations. STARS-c generates maps of chl-*a* concentrations and provides chl-*a* concentrations over time. Historically, remote sensing data use required significant computational resources because of the data size. STARS-c leverages GEE which performs computations remotely and requires only a web browser. STARS-c can help manage water resources in a sustainable manner.

This paper primarily addresses UN Sustainability Goal 6: Clean Water and Sanitation, and Goals 11, 13, and 14: Sustainable Cities and Communities, Responsible Consumption and Production, Climate Action, and Life Below Water, respectively. STARS-c provides a long-term time history of water quality, including spatially distributed concentration maps which can be used to better manage water resources and waste water treatment (G6), provide input to help design and plan sustainable cities by identifying impacted water bodies and trends (G11), help locate water and waste water treatment plants and evaluate waste water treatment effectiveness over the last 40 years (G13), provide data on long-term trends to help characterize and evaluate climate impacts and potential mitigation strategies (G13), and identify both sustainable and impacted water bodies to help address issues in life below water (G14).

Key words: chlorophyl concentration, water quality trends, remote sensing

Purpose

Eutrophication of Freshwater Systems

Eutrophication of surface waterbodies, which is defined as the excessive growth of algae and other organisms due to high nutrient levels in the water column (Khan and Ansari, 2005), is a prevalent, serious, and growing problem around the world. The effects of eutrophication, which include the growth of cyanotoxin-producing Harmful Algal Blooms (HABs) (Anderson et al., 2002), cause extensive environmental damage (Christoffersen, 1996, Paerl and Otten, 2013) in addition to economic losses in the form of increased drinking water treatment costs, spending on recovery of threatened and endangered species, diminished recreational use, and decreased value of waterfront real estate (Dodds et al., 2009). Eutrophication-related HABs also cause adverse

health effects for humans who are exposed or drink the water containing the toxins (Falconer, 1999, Chorus and Welker, 2021). It is known that human activities contribute to eutrophication of surface waters by influencing the loading of growth-limiting nutrients to aquatic systems (Caraco, 1995) and that increases in human population density and land use alterations are associated with eutrophication of nearby waterbodies (Smith, 2003).

Eutrophication of waterbodies due to human activity is predicted to continue to increase (Bennett et al., 2001, Paerl and Huisman, 2009), so there is great need for tools that will help water managers evaluate strategies for preventing and remediating eutrophication. Some success at preventing and remediating eutrophication has been achieved through methods such as controlling sources of nutrient pollution and conducting biological restoration efforts (Zhang et al., 2020, Stuart, 2001); however, in order to predict the success of these strategies, which are often costly and difficult to implement, water managers must understand the nature and cause of the eutrophication (Carpenter et al., 1999). Examining historical data can provide significant insight into this question, but very few lakes and reservoirs around the world have any kind of historical water sampling data, let alone the comprehensive dataset that would be required to effectively analyze spatial and temporal trends in water quality.

Water Quality Analysis using Remotely-sensed Data

Using remotely-sensed data addresses the lack of historical and spatially comprehensive water sampling data—researchers have successfully used satellite data to examine water quality since the early 1970's (Strong, 1974, Klemas, 2012, Richardson, 1996, Stumpf, 2001). The NASA Landsat satellite series, which has been generating useful data since 1984, collect multispectral images which can be used to evaluate plant growth, water quality, and other parameters (Brezonik et al., 2005, Hansen et al., 2015, Strong, 1974). Landsat images can be used to calculate estimates of chlorophyll-*a* (chl-*a*), a plant pigment often used as an index for algal biomass (Shi et al., 2019). Despite the ecological differences in freshwater systems around the world, they respond in similar ways to excess nutrient availability in the water column—specifically, with a marked increase in algal biomass and similar changes in species composition (Smith, 2003). Because of this, measuring changes in algal biomass (using chl-*a* as an index) is an effective method of studying trends in eutrophication for lakes and reservoirs anywhere in the world. Until recently, such analyses using remotely-sensed data were difficult and time consuming due to the large amount of data and the processing required to get the data in a usable form. The release of the Google Earth Engine (GEE) platform greatly simplified the work of retrieving, processing, and analyzing remotely sensed data, and makes it possible to analyze satellite data with just a web browser (Hansen, 2015).

STARS-c: A Tool for Analyzing Algal Blooms with Landsat Data

Building off of prior research using GEE to conduct analyses of algal bloom trends with Landsat data (Hansen et al., 2013) (Hansen et al., 2015) (Tate, 2019) (Cardall et al., 2021), we developed the SpatioTemporal Analysis of Remotely Sensed chl-*a* tool (STARS-c) to further simplify the process of conducting a remote sensing study and make it possible for anyone with an internet

connection to outline any water body, select a time range, and use either a general or user-provided model to estimate chl-*a* concentrations for that water body from 1984 to the present. STARS-c can generate stand-alone maps of chl-*a* concentrations for selected dates and provide data on long-term time series of chl-*a* concentrations. It excludes "low quality" pixels, such as pixels with clouds or dry land, from the analysis. Data generated with STARS-c characterize the spatial and temporal patterns in algal biomass of the selected water body and help water managers evaluate the success of current and proposed strategies for preventing and mitigating eutrophication by helping to identify nutrient sources, regions most impacted by nutrient inflows, and other conditions affecting the water body.

Methods

Google Colabatory and Earth Engine

Google Colabatory (Colab) is a service provided by Google Research (https://colab.research.google.com/). It allows developers and users to write and use python code in a web browser. Colab requires no configuration of the user's computers, provides access to GPUs free of charge, and allows easy sharing of any developed code. Colab uses a "notebook" concept that includes both code and text cells to provide additional context for the code. GEE is integrated into Colab and, for most GEE computational requirements, the data and the code execute on Google servers and GPU units. This is important because of the large size and intensive computation needs of remoting sensing data. For long-term analysis, such as that provided by STARS-c, datasets cover over 40 years and can use anywhere between 100s of gigabytes to more than a terabyte of data. This processing occurs on GEE servers, accessed using Colab notebook. GEE resources can be accessed with python code outside of a Colab notebook, but the notebook framework provides a browser-based tool that can be used by water managers and researchers with limited access to advanced computer resources.

STARS-c requires several python libraries. When a Colab notebook such as STARS-c is opened, it creates a new virtual environment on the Google servers. We install the required libraries on this environment in the first few cells of the notebook.

To access GEE computational resources and data, a STARS-c user must have a Google Earth Engine account, and before STARS-c can be run, the notebook the user must authorize the notebook for their account. A notebook cell starts the authorization process, then provides a link for the STARS-c user to authorize the notebook to use their GEE account. Once the required libraries are downloaded and the account authorized, STARS-c is ready for use.

Selecting a Water Body

The first tasks in STARS-c requires the user to select a waterbody for analysis. STARS-c displays a global map where the user can zoom to a waterbody. The user then uses the drawing tools (shown on the left side of **Figure 1** to draw a rectangle or other polygon around the waterbody.

This area will be used for all the subsequent analysis. The analysis will include any waterbodies included in the outline.

Error! Reference source not found. shows an example where a user has selected Lake Taihu, in China, using a rectangular area. As shown by the map, this rectangle includes several small water bodies in the areas near Taihu Lake, these will be included in subsequent analysis. A user can draw a closed polygon to select a more detailed area of interest. In this way, the user can exclude surrounding wetlands from analysis or select a smaller area of the lake to analyze separately from the lake as a whole.

After a user has created the geometry to select a waterbody, the notebook cell after the map (**Figure 1**) creates a variable that contains this geometry to define the area for all subsequent processing. If a user wants to analyze a different area or refine the geometry, they should re-run the cell that generates the map, re-draw the polygon, then run the cell that obtains the geometry.
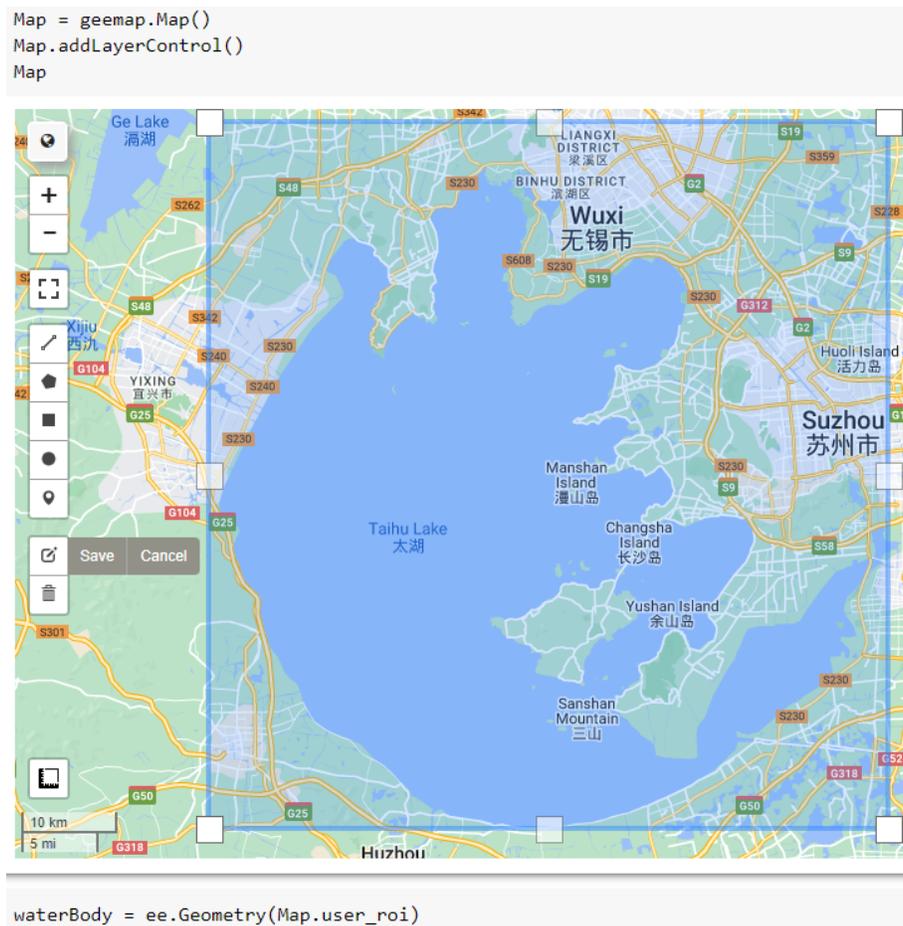
Figure 1. A cell showing the cell that displays a map, the cell displaying the map, the geometry created to select a water body, and the subsequent notebook cell that loads this geometry into a variable to define the processing area.

Satellite Data

STARS-c uses Level-2, Collection 2, Tier 1 data from Landsat satellites 5, 7, 8, and 9. The user may modify the code to use different satellite data, but it is recommended to use this data. The Landsat satellites, designed to monitor ecological conditions on the earth's surface, have a 16-day return period, 30-meter spatial resolution, a combined range of 38 years (1984-present), and spectral bands specifically designed for vegetation studies, making them ideal for developing detailed time histories of algal concentrations in waterbodies (Brezonik et al., 2005) (Masek et al., 2006) (USGS, 2016). Landsat Level-2 data include surface reflectance images, which are corrected for various sensor and atmospheric effects to best represent the spectral data that are reflected from the earth surface (https://www.usgs.gov/landsat-missions/landsat-collection-2-level-2-science-products). Level-2 data also contains pixel quality information and surface temperature data. STARS-c does not currently use surface temperature data, but we hope to include surface temperature summary statistics as an export in the future. Collection 2 images are the product of a second reprocessing effort of Landsat data and contain numerous advantages over Collection 1 data (https://www.usgs.gov/landsat-missions/landsat-collection-2). For this application, notable improvements include consistent quality assessment bands, improved radiometric calibration, and improved surface reflectance and surface temperature products. Tier 1 images are the highest-quality products from the available data. Landsat Level-2, Collection 2, Tier 1 images are readily available through Google Earth Engine.

STARS-c collects every image that contains the user-specified water body available from the Landsat archive, as shown in **Figure 2**Error! Reference source not found..

```
# Collects all Tier 1, Lav SR images of the area for each satellite
L5 = ee.ImageCollection("LANDSAT/LT05/C02/T1_L2").filterBounds(waterBody)
L7 = ee.ImageCollection("LANDSAT/LE07/C02/T1_L2").filterBounds(waterBody)
L8 = ee.ImageCollection("LANDSAT/LC08/C02/T1_L2").filterBounds(waterBody)
L9 = ee.ImageCollection("LANDSAT/LC09/C02/T1_L2").filterBounds(waterBody)
```

Figure 2. Collecting each image containing the waterbody into an image collection for each satellite (Landsats 5, 7, 8 and 9).

Data Preparation

STARS-c performs three operations to prepare the data for evaluation of chl-*a* concentration: applying the appropriate scaling factor to the bands, renaming the bands, and combining the resulting image collections into a single collection. **Figure 3** shows the code used to apply the scaling factors specified by the USGS as a function. It also shows the code used by STARS-c to apply or map this function over the different Landsat image collections. The scaling factors used for Collection 1 data are different from Collection 2 data; if the user wishes to use Collection 1 data, they will need to modify the code to apply the

correct   scaling   factors.

```python
# function to apply scaling factores
def applyScaleFactors(image):
  opticalBands = image.select('SR_B.').multiply(0.0000275).add(-0.2);
  thermalBands = image.select('ST_B.*').multiply(0.00341802).add(149.0)
  image = image.addBands(opticalBands, None, True).addBands(thermalBands, None, True)
  return image

L5Scaled = L5.map(applyScaleFactors)
L7Scaled = L7.map(applyScaleFactors)
L8Scaled = L8.map(applyScaleFactors)
L9Scaled = L9.map(applyScaleFactors)
```

Figure 3. Applying the USGS-specifed scaling factors to the bands in the image. Note that optical bands and thermal bands have different scaling factors and offsets.

Renaming the bands is not necessary, but it both simplifies computations and makes the code, especially model expressions, more readable. If the bands were not renamed, we would need different function expressions for each Landsat mission; by renaming the bands, a single expression can be used for all the Landsat missions. **Figure 4** shows the function that renames the bands along with the GEE code the applies (i.e., "maps") the function over each image collection. Landsat 5 and 7 data have different band designations from Landsat 8 and 9, so different functions are used.

```python
# function to rename Landsat 5 & 7 bands
def L57_BandRenamer(image):
  oldNames = ['SR_B1','SR_B2','SR_B3','SR_B4','SR_B5','SR_B7','QA_PIXEL','ST_B6']
  newNames = ['Blue','Green','Red','NIR','SWIR1','SWIR2','QA_PIXEL','SurfTempK']
  image = image.select(oldNames).rename(newNames)
  return image

# function to rename Landsat 8 & 9 bands
def L89_BandRenamer(image):
  oldNames = ['SR_B1','SR_B2','SR_B3','SR_B4','SR_B5','SR_B6','SR_B7','QA_PIXEL','ST_B10']
  newNames = ['UltraBlue','Blue','Green','Red','NIR','SWIR1','SWIR2','QA_PIXEL','SurfTempK']
  image = image.select(oldNames).rename(newNames)
  return image

L5Renamed = L5Scaled.map(L57_BandRenamer)
L7Renamed = L7Scaled.map(L57_BandRenamer)
L8Renamed = L8Scaled.map(L89_BandRenamer)
L9Renamed = L9Scaled.map(L89_BandRenamer)
```

Figure 4. Renaming the bands STARS-c uses in analysis. A different function is used for Landsats 5 and 7 than Landsats 8 and 9, because they have different band designations.

Merging each satellite collection into a single image collection is also not strictly necessary, but it also simplifies computation. STARS-c data exports include the name of the satellite for each image in case it is needed. The code that merges the collections is shown in **Figure 5**.

```
# Merges the band-renamed collections
# and sorts the new collection by date
renamedCollection = L5Renamed.merge(L7Renamed).merge(L8Renamed).merge(L9Renamed)
renamedCollection = renamedCollection.sort('system:time_start')
```

Figure 5. Code to merge image collections.

Image Processing

Land Masking

STARS-c estimates the chl-*a* for water pixels only. We accomplish this by first trimming the image to the boundaries of the user-drawn polygon for the water body, as shown in
Figure 6.

```
# Clips image to area polygon
image = image.clip(waterBody)
```

Figure 6. Trimming the image to the polygon drawn for the water body. All pixels inside this polygon will be evaluated for the presence of water and pixel quality, and chl-*a* concentration will be estiamted for each pixel.

STARS-c then applies a water mask using a user-specified water index. Three water indices to choose from are available in a dropdown menu, as shown in
Figure 7: the Normalized Difference Water Index (NDWI), the Modified Normalized Difference Water Index (MNDWI), Normalized Difference Vegetation Index (NDVI).

```
waterIndex:   NDWI
              NDWI
Show code     MNDWI
              NDVI
```

Figure 7. Drop-down menu to select a water index. Information about these indices is provided in a markdown cell to assist the user in selecting an index.

The user may also modify the code to specify their own water index of choice. The user must specify a threshold value to distinguish land from water. Users should modify this threshold using the visual feedback STARS-c provides to find a number that accurately represent water areas for the waterbody of interest. The water masks are sensitive to water color, turbidity, and other issues, and each water body generally requires some trial and error to select the correct threshold value. Once these parameters have been specified, STARS-c will construct a water mask, as shown in
Error! Reference source not found.. Error! Reference source not found. shows the polygon-trimmed RGB image of Lake Taihu (left), the mask that the code creates for this image (middle), and the resulting land-masked image (right).

```
# Computes water index value of each pixel
# and creates land mask
waterIndex = image.expression(waterIndexExpression,
  {
    'Green': image.select('Green'),
    'Red': image.select('Red'),
    'SWIR1': image.select('SWIR1'),
    'NIR': image.select('NIR')
  })
landMask = waterIndex.gt(waterThreshold)
image = image.updateMask(landMask)
```

Figure 8. Masking pixels that contain land. The code gives each pixel a water index based on the prodived model and masks any pixels that are not withing a specified threshold.
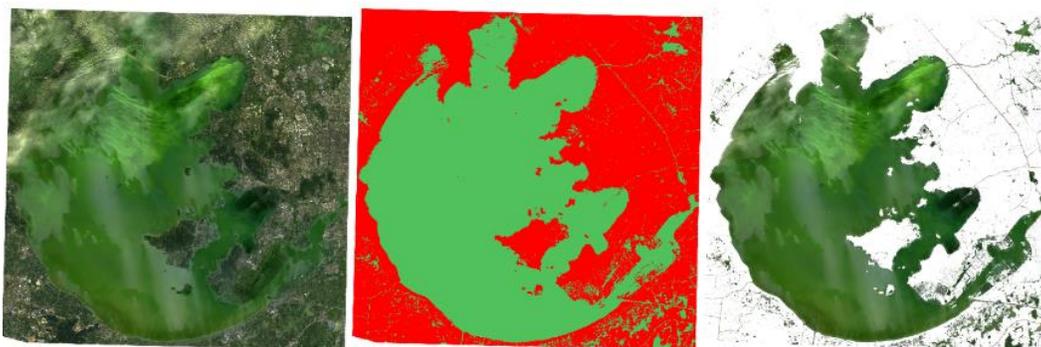
Figure 9. Land masking an image of Lake Taihu. The left panel shows the real-color image, the middle panel shows the land mask (with red indicating expelled land pixels),and the right panel shows the land-masked real-color image.

Quality Masking

STARS-c eliminates pixels containing clouds and cloud shadows from analysis by applying a cloud mask to each image. The cloud mask uses the pixel quality assessment band available in all Level 2 data. Use of this band requires bitwise operation, as shown in
Figure 10. For Collection 2 data, bit 3 of the quality band indicates the presence of a cloud and bit 4 indicates the presence of a cloud shadow. We use bit 9 to set our "cloud confidence" to high to ensure bad pixels are masked.
Figure 11 shows a land-masked RGB image of Taihu Lake (left), the cloud mask that the code creates for this image (middle), and the resulting quality-masked image (right).

```
# Pixel quality mask
qa = vizImage.select('QA_PIXEL')
# If the cloud bit (3) is set
# and the cloud confidence (9) is high
# or the cloud shadow bit is set (3)
# it's a bad pixel.
cloud = qa.bitwiseAnd(1 << 3).And(qa.bitwiseAnd(1 << 9)).Or(qa.bitwiseAnd(1 << 4))
# Remove edge pixels that don't occur in all bands
mask2 = vizImage.mask().reduce(ee.Reducer.min())
# mask image
image = image.updateMask(cloud.Not()).updateMask(mask2)
```

Figure 10. Masking low-quality pixels, including those with clouds and cloud shadows. The code can be read to say: if bits 3 and 9 are set, i.e., there are clouds and confidence in there being clouds is high, or if bit 4 is set, i.e., there is a cloud shadow, then mask the pixel.
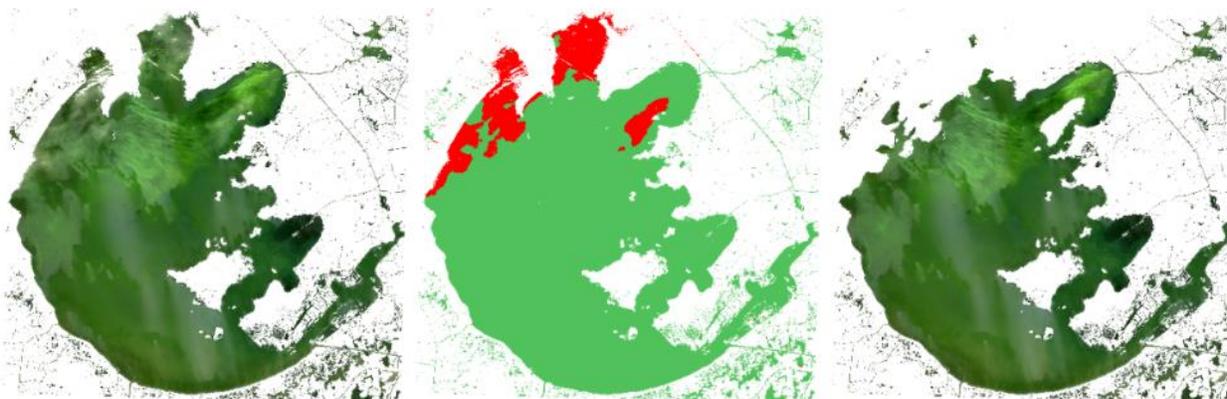
Figure 11. Quality masking an image of Lake Taihu which contains some clouds and cloud shadows. The left panel shows the real-color image, the middle panel shows the cloud mask (with the red portion inidicating the expelled pixels), and the right panel shows the masked real-color image.

Chl-a Concentration Computation

At this stage, the images are ready to be analyzed for chl-*a*. STARS-c presents seven chl-*a* models from the literature in a drop-down menu that can be selected by the user, as shown in **Figure 12**. This drop-down menu is accompanied by a description of each of the models, as well as the model.

The models presented by STARS-c were developed for specific water bodies with their own algal growth limitations and other unique properties. We strongly recommended using or developing a chl-*a* model for the waterbody being analyzed using in-situ data, if possible. If such a model is available, the user can modify the code to use this model. This modification is relatively simple, as users can use the existing models as examples of how to add a new model (or replace an existing model). In many cases, the site-specific model will use the form of one of the provided models and only the coefficients will need to be changed.

Figure 12. Dropdown menu to select a chl-*a* model. Information about these indices is provided in a markdown cell to assist the user in selecting an index.

Error! Reference source not found. shows the code which applies the model selected for analysis. This code shows the names of the Landsat bands that are used in the model definition. We renamed the Landsat bands to common terms because this allows a single model expression to work across the different Landsat missions, which use different band designations.

```
# Computes chl-a concentration of each pixel
chlaImage = image.expression(chlaExpression,
  {
    'Blue': image.select('Blue'),
    'Green': image.select('Green'),
    'Red': image.select('Red'),
    'SWIR1': image.select('SWIR1'),
    'SWIR2': image.select('SWIR2'),
    'NIR': image.select('NIR')
  })
```

Figure 13. Applying the specified chl-*a* model to the image. All pixels determined to be water and free of clouds and cloud shadows will be evaluated for chl-*a*.

Metadata

For analysis of long-term chl-a trends, it is useful to retain metadata from the original satellite data. STARS-c retains and exports the date of the satellite image in two different formats, "MM/dd/yyyy", and "MMM dd yyyy hh:mm:ss", this are included in columns titled "Date" and "Date2" in the exported .csv file, respectively. In addition, STARS-c retains and exports the satellite name (e.g., Landsat5 or Landsat8) as a column in the exported .csv file. These data can help analysis completed outside STARS-c using the exported data.

Summary Statistics

STARS-c computes and can export statistical parameters for the waterbody selected for analysis. This is done for each image in the collection that contains data--some images contain no data due to complete cloud cover or other issues. The statistics are for the selected waterbody and only include pixels with viable data but will include any water pixels in the defined geometry, even if they are not part of the main waterbody. Pixels are not included in the statistics if they are

masked; for example, pixels can be masked because they contain land or clouds or are poor quality. The number of pixels included as data for each image is also computed and exported. The number can change for each image, as each image can contain a different number of useful pixels because of clouds or changing water levels that expose or cover shoreline.

STARS-c computes the count (number of pixels used in the analysis for that image or time step), median, mean, standard deviation, minimum, maximum, skew, and the 5th, 10th, 25th, 50th, 75^{Th}, 90th, and 95th, percentiles. The 50th percentile is the same as the median, so it is reported twice. STARS-c computes these data for each image in the collection. When STARS-c exports the data, the two date formats and the satellite which collected the image are also added as columns in the .csv file.

Visualization

STARS-c provides a set of cells to visualize the computed chl-a concentrations as spatial images, and for various masks and other information used in the computation. In addition to an interactive map cell where users can toggle various layers on-and-off, change color ranges, and modify other display characteristics, this cell provides links to images of the various masks and the final chl-a concentration image. These links open the images in individual browser tabs where the image can be viewed, copied, or downloaded.

Exports

STARS-c can export the long-term time series data to a .csv file on the user's Google Drive. The .csv files will be placed in a folder titled "STARS-c_Output". STARS-c exports the data using an asynchronous background process. Depending on the size of the area and the number of images in the collection, this can take some time--it usually completes in about 30 minutes to an hour, though often sooner. STARS-c is not able to show the status of these exports. However, if the user logs into the GEE code editor (which only supports JavaScript), there is a "task" tab which the user can use to determine the status of the export.

The relatively short time required to compute and export these statistics shows the power of using the GEE servers for processing. This process analyzes over 1,000 images for most areas, computes multiple statistics using each pixel in the image, and exports the resulting data. The data are in a format that users, managers, and researchers are familiar with: a .csv file with dates and values in sequence. These data can then be used for further analysis.

Findings

STARS-c Output Examples

The following examples demonstrate the basic use of STARS-c in generating visualizations of chl-a on a waterbody and analyzing time series data. For the time series plots, the median chl-a measurement for the whole lake in each image was used; and data points above a chl-a

concentration of 300 μg/L were excluded because they are likely bad values that the masks failed to exclude and make graphs unreadable. When analyzing data, STARS-c users should pick their own outlier exclusion thresholds based on the data generated and general knowledge of typical algae concentrations in the waterbody of interest. These examples focus on turbid lakes because the high-turbidity model included in STARS-c has been well-studied.

Lake Taihu

Lake Taihu is a large, turbid, very shallow lake in the Jiangsu province in China. It is an important water resource for the area, but has attracted concern due to intense eutrophication caused by aquaculture in the eastern part of the lake, urban pollutants, and other human-caused nutrient loading.

Figure **14** shows an image from August of 2016 masked and visualized with STARS-c. From these images, it's clear that on that particular day algal blooms were most intense on the northwest side of the lake, with lower concentrations towards the middle.
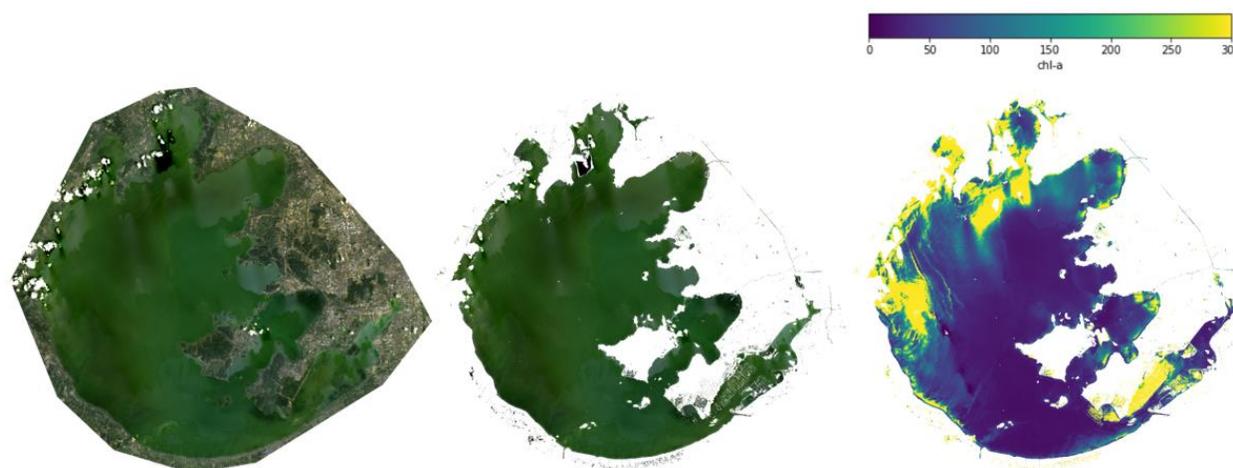
Figure 14. The polygon-cut real-color image (left), land- and quality-masked image, and chl-*a* image of Lake Taihu in August 2016.
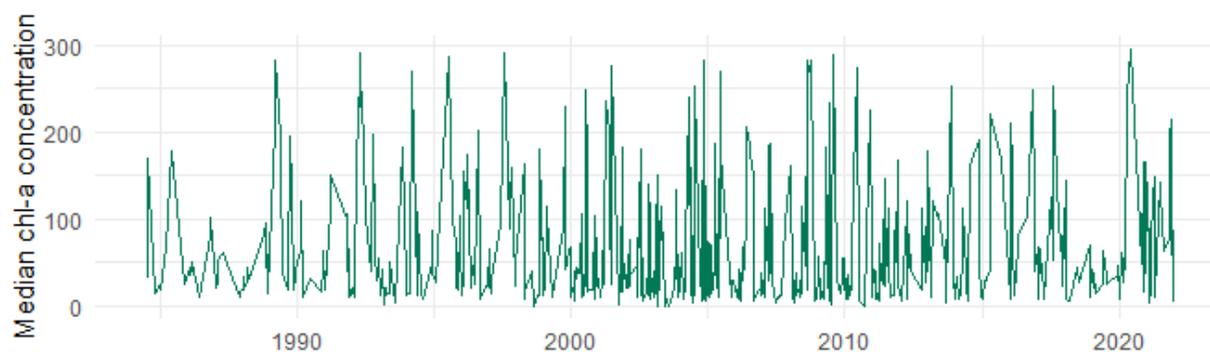
Figure 15 is an example of the type of visualization that can be created with the data exported by STARS-c. The plot shows the median chl-a concentration in every image of Lake Taihu collected

by Landsat over the ~40-year period analyzed. The data are sparser at the beginning of the series because there was only one satellite collecting data—Landsat 5—but as successive missions were launched and old satellites remained operational, the density of the data points increased, leading to more frequent data for later years.
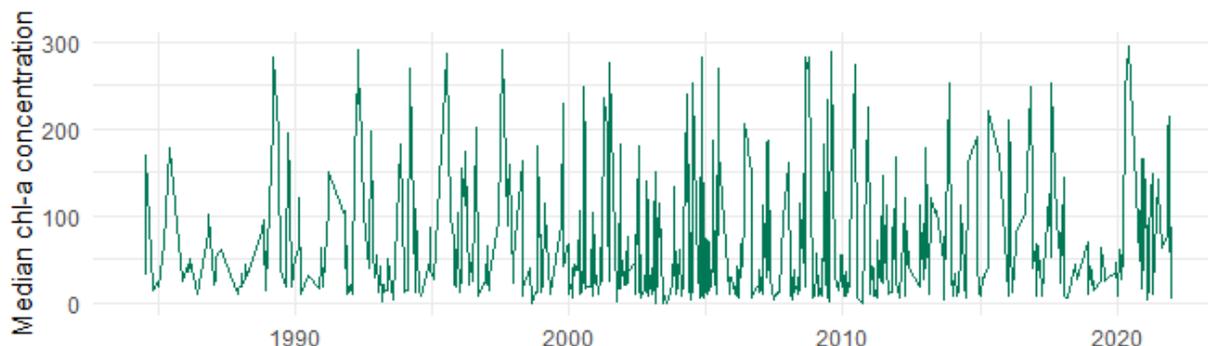
Figure 15. Time series history of median chl-a in Lake Taihu from 1985 to 2022

Malheur Lake

Malheur Lake is another large, shallow lake with high-turbidity water located in eastern Oregon, United States. It has also faced challenges associated with eutrophication and high nutrient loads, and in the Landsat image from August 2013 visualized in
Figure **16**, it is clear that algae biomass in the lake can be quite large.

Figure 16. The polygon-cut RGB image (left), land- and quality-masked image, and chl-*a* image of Malheur Lake in August 2013.

The time series plot for Malheur Lake, shown in
Figure **17**, shows that the algae concentrations are highly variable but follow a distinct seasonal trend with higher concentrations in the summer, which is to be expected.
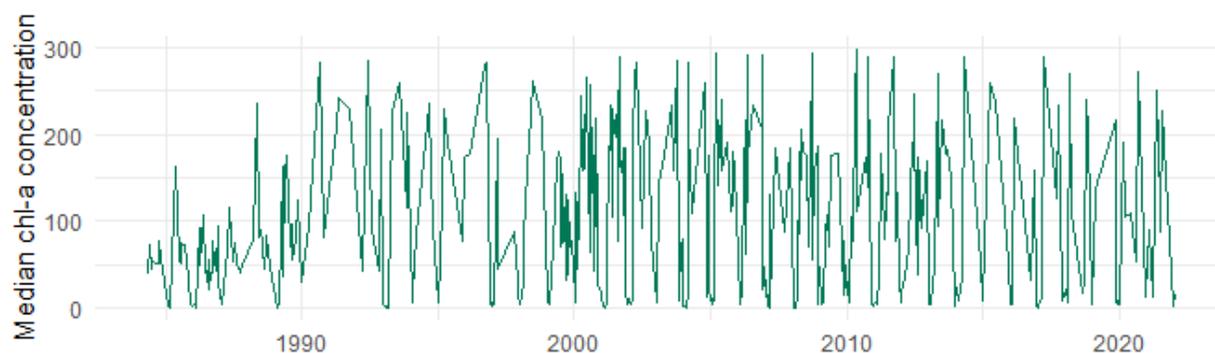
Figure 17. A time-series plot of the average chl-a concentration for Malheur Lake over the nearly 40-year period of Landsat data.

Utah Lake

Utah Lake, in the northern part of Utah, United States, shares many characteristics with Taihu Lake and Malheur lake, and is also the subject of much concern over eutrophication. We have conducted remote-sensing and other types of research on this lake prior to the development STARS-c, and with this background knowledge of the lake's ecology and hydrology, the visualizations and time series data generated with STARS-c become much more meaningful and useful. For example, we know that the bay on the east side of the lake (Provo Bay) is very shallow and typically has lower-turbidity water than the rest of the lake. In an image from July of 1986, shown in

Figure **18**, algal concentrations in Provo Bay are higher than much of the lake, so this combined with the likelihood of that water being clearer has interesting implications for our understanding of what factors contribute to algal growth in Utah Lake. In addition, we know that the eastern shore of the lake is significantly more developed than the western shore, so we can examine spatial distributions of chl-*a* within that context to provide further insights into how shoreline development might impact algal blooms.
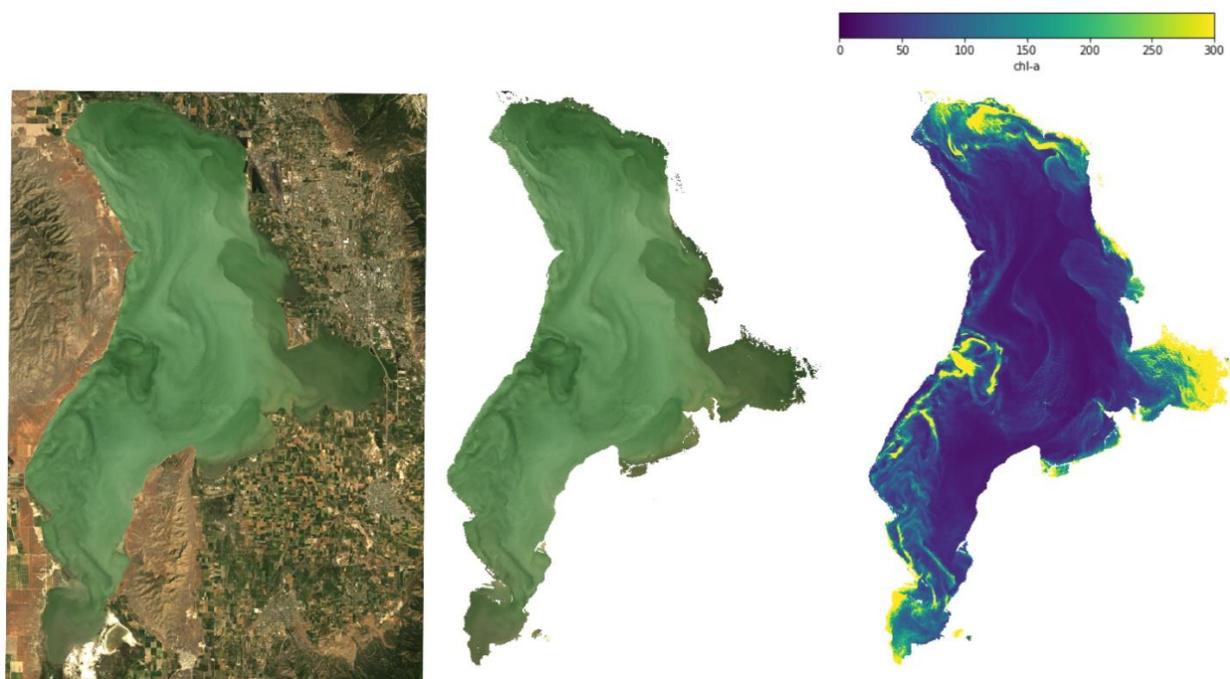
Figure 18. The polygon-cut image (left), land- and quality-masked image, and chl-*a* image of Utah Lake in July 1986. Grey swirls are regions with high suspended sediemnts, the darker green color correlates with higher chl-a concentrations.

Figure 19 shows the time series plot of chl-a concentrations on Utah Lake. In this case also, prior knowledge of the lake allows us to analyze the data more rigorously. For example, we can examine how past and current river delta restoration projects on Utah Lake tributaries may have impacted algal concentrations in the lake at the outlet of those tributaries, or determine to what extent algal concentrations correlate with the dramatic population growth on the eastern shore that has occurred over the last 40 years.
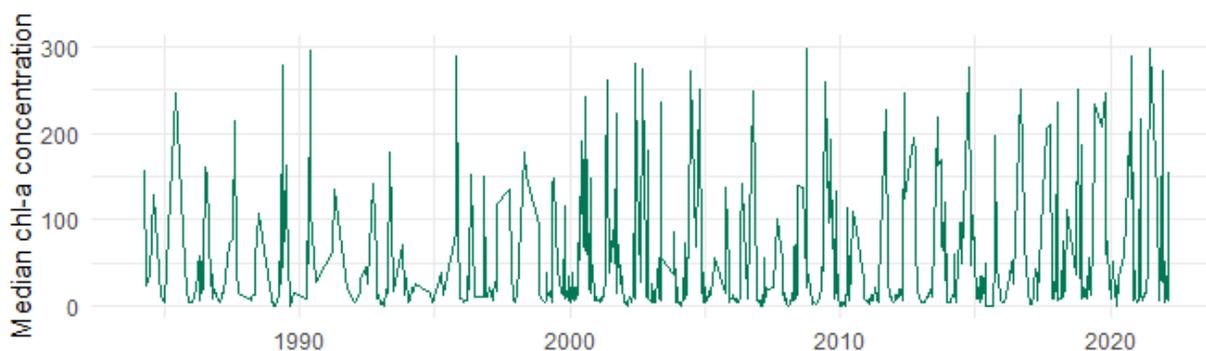
Figure 19. A time-series plot of the average chl-a concentration for Utah Lake over the nearly 40-year period of Landsat data.

Deer Creek Reservoir

This example was included to show how STARS-c handles a waterbody with clear water and typically low concentrations of algae. Deer Creek Reservoir, in the northern part of Utah, United States, is very deep and clear, with much lower temperatures than the other three lakes; though in recent years the Utah State Department of Water Quality has observed harmful algal blooms and there is some concern over the ecological health of the reservoir.

Figure **20** is a visualization of a Landsat image from July of 1986 and shows that concentrations of algae in Deer Creek on that day were extremely low. The model used to calculate chl-a concentrations for this reservoir was the Clear Whole Season model included in STARS-c, which was developed specifically for Deer Creek using water sampling data, so this model provides more accurates result for this lake than the Turbid Whole Season model used for the other three lakes would.
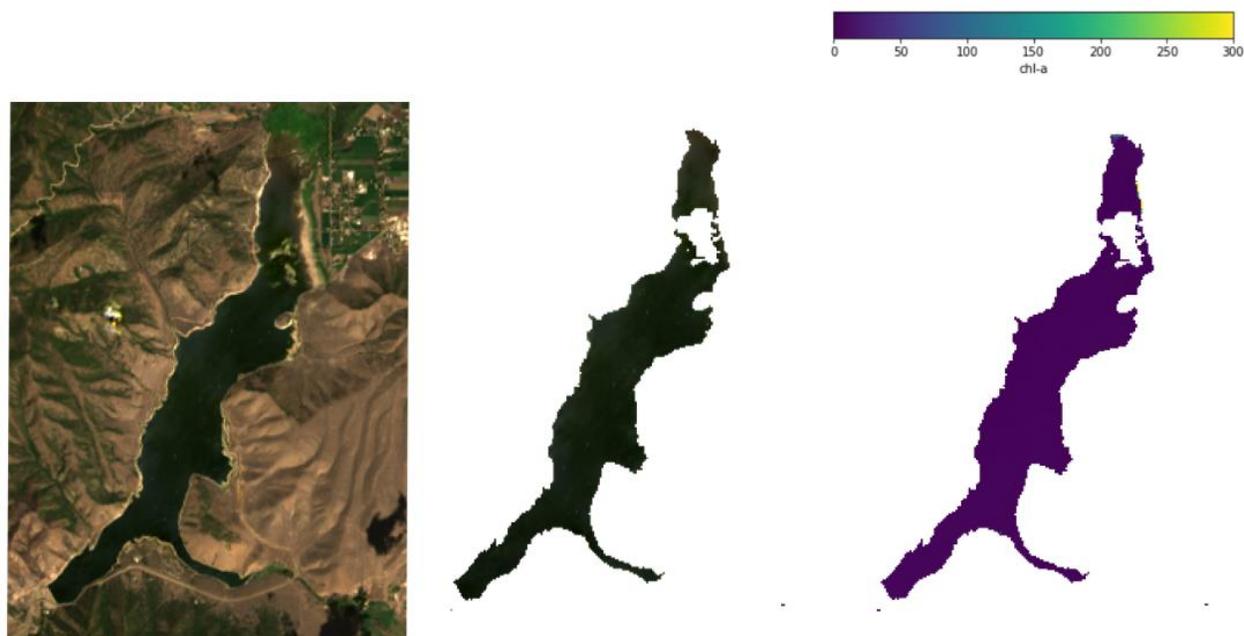
Figure 20. The polygon-cut RGB image (left), land- and quality-masked image, and chl-*a* image of Deer Creek in July 1986. There is a small cloud and cloud shadow in the northern part of the image that results in the area being masked. Chl-a concentrations in Deer Creek are significantly lower than Utah Lake, which is approximately 25 km downstream.

Figure 21, the time series visualization of median chl-a data on Deer Creek Reservoir, looks quite different from the time series of the more turbid, eutrophic lakes. In this case, it appears that algal concentrations in Deer Creek are typically very low, with the occasional large spike in concentration representing a large bloom. From this graph, it appears that there could be a trend

of increasing frequency for these blooms, however, this may be due to the higher frequency of data in late years, which is the result of more satellites collecting data.
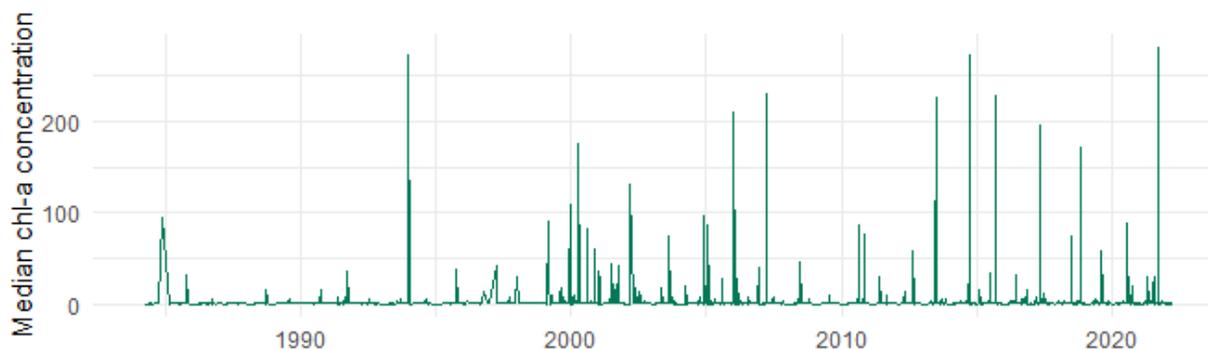
Figure 21. A time-series plot of the average chl-a concentration for Deer Creek reservoir over the nearly 40-year period of Landsat data.

Research Limitations

While STARS-c outputs data as chl-a concentrations, these measurements may not be highly accurate for any given water body due to the limitations inherent in remote sensing. The data can still be useful, though, because the change in concentrations between images or over the long-term should be relatively accurate, and therefore can be used effectively to analyze spatial and temporal trends. To obtain accurate concentrations, the user would need to use in-situ data from the water body to fit a chl-a model specific to that lake or reservoir. While STARS-c provides a long-term history of chl-a concentrations, it is important to understand the dynamics and processes of the specific waterbody of interest. For example, if there is a large nutrient source, such as a waste-water treatment plant, the integrated statistics would not necessarily show that impact. Individual images could show some impact, depending on the size and circulation patterns of the lake, but would not include the long-term trend. Users can place smaller geometric regions around a lake, isolating various areas, and generate the long-term statistics for that area. These segments can then be used to compare different regions of the lake or to compare a small area to the larger lake as a whole in order to provide better insight into spatial variation.

Originality and STARS-c Value

STARS-c allows water managers and researchers to visualize, evaluate, and analyze long-term trends in chl-*a* concentrations in waterbodies across globe. It runs in a browser and uses cloud resources, meaning that STARS-c is not resource constrained and requires minimal computational infrastructure. Access to STARS-c is available upon request.

References

Anderson, D. M., Glibert, P. M. & Burkholder, J. M. 2002. Harmful algal blooms and eutrophication: Nutrient sources, composition, and consequences. *Estuaries,* 25**,** 704-726.

Bennett, E. M., Carpenter, S. R. & Caraco, N. F. 2001. Human Impact on Erodable Phosphorus and Eutrophication: A Global Perspective: Increasing accumulation of phosphorus in soil threatens rivers, lakes, and coastal oceans with eutrophication. *BioScience,* 51**,** 227-234.

Brezonik, P., Menken, K. D. & Bauer, M. 2005. Landsat-based Remote Sensing of Lake Water Quality Characteristics, Including Chlorophyll and Colored Dissolved Organic Matter (CDOM). *Lake and Reservoir Management,* 21**,** 373-382.

Caraco, N. F. 1995. Influence of human populations on phosphorus transfers to aquatic systems: a regional scale study using large rivers. *Scope-Scientific Committee On Problems Of The Environment International Council Of Scientific Unions,* 54**,** 235-244.

Cardall, A., Tanner, K. B. & Williams, G. P. 2021. Google Earth Engine Tools for Long-Term Spatiotemporal Monitoring of Chlorophyll-a Concentrations. *BYU ScholarsArchive,* 7**,** 4.

Carpenter, S. R., Ludwig, D. & Brock, W. A. 1999. Management OF Euthrophication For Lakes Subject To Potentially Irreversible Change. *Ecological Applications,* 9**,** 751-771.

Chorus, I. & Welker, M. 2021. *Toxic cyanobacteria in water: a guide to their public health consequences, monitoring and management*, Taylor & Francis.

Christoffersen, K. 1996. Ecological implications of cyanobacterial toxins in aquatic food webs. *Phycologia,* 35**,** 42-50.

Dodds, W. K., Bouska, W. W., Eitzmann, J. L., Pilger, T. J., Pitts, K. L., Riley, A. J., Schloesser, J. T. & Thornbrugh, D. J. 2009. Eutrophication of U.S. Freshwaters: Analysis of Potential Economic Damages. *Environmental Science & Technology,* 43**,** 12-19.

Falconer, I. R. 1999. An Overview of problems caused by toxic blue–green algae (cyanobacteria) in drinking and recreational water. *Environmental Toxicology,* 14**,** 5-12.

Hansen, C. 2015. *Google Earth Engine as a Platform for Making Remote Sensing of Water Resources a Reality for Monitoring Inland Waters*.

Hansen, C., Swain, N., Munson, K., Adjei, Z., Williams, G. P. & Miller, W. 2013. Development of Sub-Seasonal Remote Sensing Chlorophyll-A Detection Models. *American Journal of Plant Sciences,* 04**,** 21-26.

Hansen, C. H., Williams, G. P., Adjei, Z., Barlow, A., Nelson, E. J. & Miller, A. W. 2015. Reservoir water quality monitoring using remote sensing with seasonal models: case study of five central-Utah reservoirs. *Lake and Reservoir Management,* 31**,** 225-240.

Khan, F. A. & Ansari, A. A. 2005. Eutrophication: An ecological vision. *The Botanical Review,* 71**,** 449-482.

Klemas, V. 2012. Remote sensing of algal blooms: an overview with case studies. *Journal of Coastal Research,* 28**,** 34-43.

Masek, J. G., Vermote, E. F., Saleous, N. E., Wolfe, R., Hall, F. G., Huemmrich, K. F., Feng, G., Kutler, J. & Teng-Kui, L. 2006. A Landsat surface reflectance dataset for North America, 1990-2000. *IEEE Geoscience and Remote Sensing Letters,* 3**,** 68-72.

Paerl, H. W. & Huisman, J. 2009. Climate change: a catalyst for global expansion of harmful cyanobacterial blooms. *Environmental Microbiology Reports,* 1**,** 27-37.

Paerl, H. W. & Otten, T. G. 2013. Harmful Cyanobacterial Blooms: Causes, Consequences, and Controls. *Microbial Ecology,* 65**,** 995-1010.

Richardson, L. L. 1996. Remote sensing of algal bloom dynamics. *BioScience,* 46**,** 492-501.

Shi, K., Zhang, Y., Qin, B. & Zhou, B. 2019. Remote sensing of cyanobacterial blooms in inland waters: present knowledge and future challenges. *Science Bulletin,* 64**,** 1540-1556.

Smith, V. H. 2003. Eutrophication of freshwater and coastal marine ecosystems a global problem. *Environmental Science and Pollution Research,* 10**,** 126-139.

Strong, A. E. 1974. Remote sensing of algal blooms by aircraft and satellite in Lake Erie and Utah Lake. *Remote Sensing of Environment,* 3**,** 99-107.

Stuart, A. L. 2001. Life after death in lake erie: nutrient controls drive fish species richness, rehabilitation. *Ecological applications,* v. 11**,** pp. 731-746-2001 v.11 no.3.

Stumpf, R. P. 2001. Applications of satellite ocean color sensors for monitoring and predicting harmful algal blooms. *Human and Ecological Risk Assessment: An International Journal,* 7**,** 1363-1368.

Tate, R. S. 2019. Landsat Collections Reveal Long-Term Algal Bloom Hot Spots of Utah Lake.

Usgs 2016. Landsat—Earth observation satellites. *Fact Sheet.* Version 1.0: Originally posted November 25, 2015; Version 1.1: August 22, 2016; Version 1.2: April 8, 2020 ed. Reston, VA.

Zhang, Y., Luo, P. P., Zhao, S. F., Kang, S. X., Wang, P. B., Zhou, M. M. & Lyu, J. 2020. Control and remediation methods for eutrophic lakes in the past 30 years. *WATER SCIENCE AND TECHNOLOGY,* 81, 1099-1113.